DOCUMENT RESUME

ED 357 033 TM 019 727

AUTHOR Moore, William P.
TITLE Preparation of Students for Testing: Teacher Differentiation of Appropriate and Inappropriate Practices.
PUB DATE Apr 93
NOTE 20p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (Atlanta, GA, April 13-15, 1993).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Classroom Techniques; Educational Practices; Elementary Education; *Elementary School Teachers; Ethics; *Paraprofessional School Personnel; Pretests Posttests; Questionnaires; Teacher Attitudes; *Teacher Behavior; Teacher Made Tests; *Test Coaching; Test Use; Urban Schools
IDENTIFIERS Large Scale Programs; Teacher Assessment Prep Practices Questionnaire

ABSTRACT

This paper studied whether or not elementary school classroom teachers in a large urban midwestern school district were able to distinguish appropriate from inappropriate testing practices in a large-scale mandated program. Fifty of 62 teachers and paraprofessionals in 2 elementary schools completed the Teacher Assessment Preparation Practices Questionnaire (TAPQ), which explored 40 specific testing behaviors of teachers from pretesting to posttesting. Respondents rated each teacher behavior regarding testing for acceptability. Participants distinguished appropriate testing behaviors, but did not demonstrate the expected capability when rating the behaviors. Less than half of the inappropriate behaviors were correctly identified. Those that were characterized as inappropriate had the largest standard errors and variability indices, indicative of disagreement among participants about the appropriateness of these practices. Teachers and paraprofessionals responded in similar ways, demonstrating similar levels of understanding of testing practice. Findings support other research results that have suggested that classroom educators are not prepared to implement appropriate and acceptable test preparation and test administration practices. Recommendations for improvement are included. One figure illustrates the discussion, and four tables summarize responses to the questionnaire items. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED357033

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

William P. Moore

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Preparation of Students for Testing:

Teacher Differentiation of Appropriate and Inappropriate Practices

William P. Moore

The University of Kansas

For correspondence and preprints:

William P. Moore, Ph. D.
Senior Research Coordinator
Assistant Professor
The University of Kansas, Medical Center
4030 Robinson Hall
3901 Rainbow Blvd
Kansas City, KS 66160
(913) 588-4703

Paper presented at the annual meeting of the National Council on Measurement in Education, Atlanta, GA, April, 1993.

TM019727

BEST COPY AVAILABLE

Preparation of Students for Testing: Teacher Differentiation

of Appropriate and Inappropriate Practices

YOUR STUDENTS' SCORES ARE TOO LOW!. . . What are teachers to do? Many public school educators are unprepared to implement an effective, appropriate test preparation program designed to improve student achievement test performance. While not a recent revelation, evidence points to a lack of testing and assessment traning for pre- and in-service teachers. The paucity of training has been noted by Schafer and Lissitz (1989) who found that roughly 50% of the teacher training programs in the United States require measurement coursework for teacher certification. Stiggins, Conklin and Faires (1989), in a review of assessment curriculum in 27 undergraduate and graduate teacher training programs, found that less than half even provided assessment instruction and only 6 required the training for graduation. Others (Gullickson, 1986; Stiggins, 1987) have found in-service teachers who did receive assessment training to be practicing forms of assessment not covered during their training and have reported that their training was not relevant to their assessment information needs.

While teachers appear to be suffering from a lack of training to effect improved test performance through appropriate and ethical means, there exists an increasing proliferation of mandated testing programs and decisions based on these programs. "Standardized tests are used increasingly in evaluating the quality of the local schools. This places pressure on the administrators and teachers to engage in activities that are intended to increase students' scores" (Mehrens & Kaminski, 1989). Current speculation regarding testing behavior suggests that teachers, sensing the importance of performance on achievement tests to students and themselves, heighten efforts to demonstrate increased test scores. The assumption that pressure forces teachers to take steps to enhance test performance, sometimes appropriate and other times inappropriate, has not been fully investigated. Central to this assumption is test-stakes (Corbett & Wilson, 1988). Research suggests that a test's stakes are related to the importance of the decisions being made as a result of test performance and that decisions that affect future access can lead to pressure to perform well on the test. While this causal hypothesis seems logical, some evidence does exist suggesting that pressure is not the only factor motivating teacher testing behavior. Moore (1992), in a study of ITBS testing, found zero-order correlations suggesting a non-significant relationship between perceived pressure

and engagement in inappropriate practices. Alternately, the perceived value and derived benefits of the testing program were found to be significantly inversely related to engagement in inappropriate testing practices. In both high stakes and low stakes testing settings, teachers are confronted with administrator pressure to increase test score gains "using any means available" (Moore, 1992). An apparent inattention by administrators and teachers to violations of the standardization principles of preparation, administration, and norming associated with standardized testing, is not surprising given an educational climate focused on test score improvement and not necessarily improved instructional practices and learning skills. Consequently, teachers in many different testing settings have reported little confidence in the results of standardized achievement tests or the value of test score information (Haas, Haladyna & Nolen, 1989; Moore, 1992; Rottenberg & Smith, 1990). The scenario described above makes a broad assumption that teachers willingly violate testing principles to reach the goal of test score gains. A more likely explanation may exist: teachers are not trained to recognize or understand the measurement and standardization principles of testing.

Many questions regarding test preparation are central to understanding the forces that may lead teachers to engage in inappropriate testing practices. Others (Fish & Allard, 1990; Glasnapp, Poggio, & Miller, 1991; LeMahieu & Wallace, 1986; Madaus, 1987; Mehrens & Kaminski, 1989) have explored the impact of mandated testing programs and pose a number of valuable questions. However, few have asked: are teachers able to discriminate between what is and is not appropriate testing-related practices for a given assessment situation? Research reported eight years ago suggested that teachers were not able to distinguish between "cheating" practices and acceptable practices (Gonzalez, 1985). More recently, Popham (1991), examined five broad forms of preparation (previous-form preparation, current-form preparation, generalized test-taking preparation, same format preparation, and varied-format preparation). His results indicated substantial variation among educators regarding the appropriateness of these types of activities.

The current study had the overall goal of identifying gaps in teacher knowledge of standardized testing practices. The primary objective of this pilot study was to determine whether elementary classroom teachers in one urban midwestern school district were able to distinguish appropriate from inappropriate

testing behavior within the context of a large-scale mandated testing program. A second objective was to identify the perceived similarities of testing behaviors among teachers. The study tested the following hypotheses:

Hypothesis One: Educators in this sample will not be able to accurately identify appropriate and inappropriate testing practices within the context of a standardized, norm-referenced, mandated testing program.

Hypothesis Two: Educators in this sample will perceive appropriate and inappropriate testing behaviors to be similar, based on factor analysis loadings, demonstrating a lack of discrimination.

Measures

Teachers were presented with a survey instrument (Teacher Assessment Preparation Practices-TAPP^{e}). A similar version (Teacher Assessment Practices Questionnaire-TAPQ) was previously utilized in a study of teacher testing practices (Moore, 1991, 1992). The TAPP^{e} contained 40 specific testing behaviors spanning pre-testing to post-testing. Respondents were asked to rate each testing behavior as an 1) "acceptable practice"; 2) "questionable but still ethical practice"; 3) "unacceptable but not outright cheating"; and 4) "unethical or cheating" within the context of ITBS testing. One sample item follows:

——**6. Use prior year test questions as practice for this year's test.**

One item (ROLE) asked participants to indicate their professional position (teacher or paraprofessional). A second item (GRADE) asked participants to indicate the grade level of students most often seen.

Reliability and validity of the TAPQ instrument were reported in prior work (Moore, 1991, 1992) and was found to have moderate to high stability coefficients for scales built with the 40 items (.4 to .9). Original instrument development for the 40 behaviors was based on conventional wisdom and information and not on pre-identified scales. Content validity for representation of the domain of testing practices was established through the input of 21 assessment directors in as many states and through a thorough review of the assessment literature. Two measurement specialists within the district examined the 40 behaviors in order to establish a categorization of appropriateness and rate each item as either appropriate or inappropriate given the intent of the ITBS and the level of generalization desired within the district.

Twenty-three items were rated as inappropriate for the ITBS and 17 items were found to be appropriate by review of these specialists.

The results of this study found internal consistency reliability for the TAPP^c 37 items with non-zero variance to be $r_\alpha = .804$ (standardized). Three items with zero variance (Q2, Q5, Q23) were not included in the reliability estimate.

Sample and Methods

Sixty-two teachers and paraprofessionals (paras) employed in two different elementary schools in one large midwestern urban school district were asked in the spring of 1992 to participate in this pilot study. The teachers and paras were attending a district staff development session regarding ITBS testing. The session occurred approximately four weeks before testing.

Prior to staff development presentations the participants were asked to complete the TAPP^c referencing their perceptions of ITBS testing. Because of the sensitivity of the topic and the heightened awareness of testing issues in the district, as well as the setting in which participants were asked to respond, no demographic information was obtained other than grade level taught (K through 5) and role (teacher or paraprofessional). The participants were directed to not discuss their perceptions with others in the session until all surveys had been completed and collected. While this limits the generalizability of the results, this pilot study will suggest if teachers in this midwestern urban district are able to discriminate between testing practices.

Results

The inclusion of paraprofessionals in the sample and in analyses provided a convenient comparison group of instructional participants who have less formal training in education. Of the 62 participants, 50 returned a completed survey (81% response rate: 77% teachers (42), 100% (8) paraprofessionals). Fourteen kindergarten/first grade teachers and paras, 11 second grade, 10 third grade, 10 fourth grade, and 4 fifth grade teachers and paras responded to the instrument.

Accuracy of Ratings

For the 17 practices considered <u>appropriate</u> by specialist review, the mean accuracy of study participants was 15.47 practices or 91% . However, when the 23 <u>inappropriate</u> practices were rated by

study participants the mean accuracy was 10.35 practices or 45%. Of these 23 inappropriate behaviors, 13 were considered appropriate or questionable but not inappropriate by more than 50% of the participants. Of those behaviors considered appropriate by participants, which were categorized as inappropriate by specialists, many were characteristic of a measurement-driven instructional (MDI) approach or a criterion-referenced approach (e.g., prepare instructional objectives based on test items) to testing. Others were motivational in nature and could potentially place undue pressure to perform on certain segments of the student population (e.g., talk to best students and encourage them to do their best).

Seventy-five percent (30) of the 40 practices were considered appropriate testing practices by more than 50% of the participants. As such, more than half of the participants correctly identified 10 of the 23 inappropriate behaviors and all of the appropriate behaviors. Based on χ^2 estimates we expected 50% or more of the participants to correctly identify 15 of the inappropriate practices and 12 of the appropriate practices. The discrepancy was found significant (χ^2 =14.235; df=2; $p < .0001$). While able to identify appropriate practices the participants were not able to accurately identify the majority of inappropriate practices. Ratings indicated very little consensus among teachers and paras for the 40 practices (see Table 1). Inappropriate practices most often mis-classified as appropriate were: 'encourage attendance in test week and provide rewards for high attendance' (94%), 'use commerical test preparation package (Scoring High on the ITBS)' (92%), 'prepare instructional objectives based on ITBS test items' (92%), and 'take each skill tested and direct day-to-day instruction toward these skills' (86%) (see Table 2).

Insert Tables 1, 2 about here

Teacher and paraprofessional comparison. Multivariate analysis of variance was conducted to examine accuracy of ratings of testing practices. Two factors were examined: role and grade level taught. The two-way interaction of ROLE x GRADE was found non-significant (Hotellings T = .319; p = .18). Mean accuracy differences by role were non-significant (Hotellings T = .03; p = .551). Teachers were able to correctly identify 91% of the appropriate practices and paras identified 88%. Similarly, teachers

correctly identified 45% of the inappropriate practices and paras identified 46%. As such, teachers and paraprofessionals did not differ in their ability to discriminate between appropriate and inappropriate testing behaviors.

The main effect for grade level taught was non-significant as well (Hotellings T = .350; p = .13). Cell means for accuracy of inappropriate practices ranged from a low of 38% for third grade educators to a high of 51% for kindergarten/first grade educators. Mean accuracy of appropriate practices by grade level ranged from 89% for second grade educators to 94% for fourth grade educators.

Insert Figure 1 about here--% correct

Variability of Perceptions

With few exceptions, the relative magnitude of standard errors and variability indices were larger for those behaviors categorized as inappropriate (see Table 3). Behaviors considered appropriate demonstrated the least error and variability among participants. The most variable practice 'give practice questions which are directly off the current test' had a standard error of .19. Responses indicated that 30% of the sample considered this to be an appropriate practice. The least variable practices 'teacher how to answer multiple choice questions' and 'teach students how to follow test directions' had a consensus among educators with 100% acceptance. The greater variability found for inappropriate behaviors suggests that the educators in this study did not hold a consensual view regarding the appropriateness of testing practices.

Insert Table 3 about here -- variability

Perceptual Similarities of Testing Practices

To determine if perceptual similarities between appropriate and inappropriate practices could be identified, responses to the 40-item survey were submitted to exploratory factor analysis procedures. Factors composed of both types of practices would provide insight into perceptual similarities among

practices and allow for identification of the types of practices most often confused as appropriate or inappropriate.

Six factors were extracted using a common factors analysis with the highest correlation on the diagonal of the matrix. The six factors explained 47% of the common variance. Using an orthogonal rotation the six factors were: I-Inappropriate Interventions (15.5%); II-Testwiseness-Measurement-Driven-Instruction (11.2%), III- Inappropriate Item Exposure (6.3%), IV-Emphasis with Students (5.1%), V- Test-taking Skills (4.6%), and VI-Motivational/Incentives (4.3%).

Insert Table 4 about here--factor loadings

The matrix loadings support the hypothesis that educators in this sample perceived both appropriate and inappropriate practices to be similar with four of the six factors composed of both appropriate and inappropriate practices. Furthermore, an examination of the modal responses for items indicated that many inappropriate behaviors were perceived to be appropriate. For example, the modal response for the item 'teach vocabulary words found on the current test' was 1 (acceptable behavior) but this item loaded on Factor I- Inappropriate Intervention (.51). Factor II-Testwiseness/MDI demonstrates the large discrepancy in teacher understanding of appropriateness with four of seven items clearly inappropriate for a norm-referenced standardized test such as the ITBS, yet all four items were rated as appropriate by educators. For example, the practice 'Prepare instructional objectives based on ITBS test items' (q41) is a misapplied approach commonly known as: prepare test items based on instructional objectives. The reversal of this criterion-referenced approach in which instruction determines test content is now conceptualized as test content determining instructional practice. Similarly, the practice 'Focus instruction on extensive drill and practice on ITBS skills' (q28) reflects a measurement-driven approach to instruction and testing and is clearly inappropriate for a norm-referenced testing program. The fact that a dimension reflecting a measurement-driven approach was observed within the context of a norm-referenced testing program suggests considerable misunderstanding regarding appropriate testing practices.

Discussion

The emerging view of mandated testing suggests at least two potential explanations of teacher testing behavior: teachers are not trained to deliver appropriate testing preparation or to respond appropriately to calls for increased test scores; and the pressure to improve scores leads teachers to utilize inappropriate forms of preparation. A concurrent alternative view may be that teachers do not see the mandated test as a useful assessment device and do not value the test which leads to potential unethical preparation practice. The research literature provides evidence for each explanation and it is most likely that each circumstance exists.

These pilot findings suggest that the sample participants were quite capable of distinguishing appropriate testing behaviors but did not demonstrate the expected capability when rating the inappropriate behaviors. In fact, less than half of the inappropriate behaviors were correctly identified. Second, those behaviors categorized as inappropriate had the largest standard errors and variability indices indicating considerable disagreement among the participants about the appropriateness of these behaviors. Ability to differentiate among practices was found to be equally poor for both teachers and paraprofessionals with each group correctly identifying less than 50% of the inappropriate practices. Non-significant differences in accuracy between teachers and paraprofessionals suggests similar levels of understanding of testing practice irrespective of amount of pre-service and in-service training. Lastly, the results of a factor analysis indicated that many inappropriate practices loaded with appropriate practices on the extracted factors. The modal responses of inappropriate behaviors loading with appropriate behaviors suggested that participants found these practices to be appropriate and considered them perceptually similar to appropriate practices.

As such, the findings provide tentative evidence in support of other research findings suggesting that classroom educators are not prepared to implement appropriate and acceptable test preparation and test administration. Unwittingly, teachers may be engaging in inappropriate and unethical behaviors (Moore, 1992; Nolen, Haladyna & Haas, 1990) without an understanding of the appropriateness of their behaviors, the implications of violating the standardization assumption, and the intent of the testing program.

Recommendations

Schafer (1991) makes an attempt to identify the assessment skills necessary for teachers to have mastered. While he notes the importance of ethics in testing and impact of testing on *students* he leaves out any mention of impact of testing on classroom instruction, curriculum development and teacher testing practices. While it has been shown that contemporary measurement and assessment instruction focuses on item writing, test development, statistics, validity, and reliability (Gullickson & Hopkins, 1987 and Hills, 1989 in Airasian, 1991), as well as issues surrounding standardized testing, the proliferation and influence of mandated testing programs demands greater pre-service instructional attention to the "unseen" influences of testing (e.g., pressure to show gains, to teach to the test, to realign curricular objectives and test objectives, to modify instructional methods to better reflect testing scenarios, to appropriately prepare students for testing) and less instruction in the mechanics of item writing and statistics. Schafer's assessment essentials reflect only one aspect of what pre- and in-service teachers need to know: the mechanics and interpretation of assement information. Teachers need to know how to cope with the influences demanding increased student scores on tests that are increasingly being used to evaluate their own instructional performance. Within the context of pre- and in-service assessment training, the following recommendations are offered:

1. Provide pre-service teachers with a realistic view of the assessment climate in school districts through a *Psychology/Sociology of Educational Assessment Systems* curricular offering. A honest discussion of the powerful forces at work in buildings and districts could be a first step in preparing pre-service teachers for the eventual confrontation of testing vs. learning focus in educational politics.

2. Through studies such as this, staff development specialists in districts could identify misunderstanding or ignorance of assessment and preparation principles and target in-service training to address these practices in a non-threatening, informative fashion. Of course, any district-sponsored data collection would need to assure respondent annonymity.

3. National organizations such as NCME, AERA, AFT and NEA must address the consequences associated with mandated testing programs through the development of model measurement and assessment curriculum recommendations, instructional modules, annual meeting mini-courses, and

expanded discussion in the media and professional literature. Conferences exploring this topic with state department assessment specialists, local district testing directors, test developers, and university/college teacher training educators should be undertaken. Textbook authors need to become sensitive to the influences public school teachers face and attempt to develop their material with a greater understanding of the most salient assessment needs of teachers. While this pilot study provides only a tentative picture of the status of in-service teacher knowledge in one school district, the results and instrumentation used may be the foundation for exploring this problem in a broader context. The recommendations are valid even without the findings reported here.

References

Corbett, H. D. & Wilson, B. (1988). Raising the stakes in state-wide mandatory minimum competency testing. *Politics of Education Association Yearbook 1988*, 27-39.

Fish, J. & Allard, J. (1990, Arp.). *The impact of mandated standardized testing*. Paper presented at the annual meeting of the American Educational Research Association, Boston, MA.

Glasnapp, D. R., Poggio, J. P., & Miller, M. D. (1991). Impact of a "low stakes" state minimum competency testing program on policy, attitudes, and achievement. *Advances in Program Evaluation, 1B*, 101-140.

Gonzalez, M. (1985). Cheating on standardized tests: What is it? In P. Wolmut & G. Iverson (Ed.), *National Association of Test Directors 1985 Symposia* (pp. 4-16). Portland, OR: Multinomah, ESD.

Gullickson, A. R. (1986). Teacher education and teacher-perceived needs in educational measurement and evaluation. *Journal of Educational Measurement, 23(4)*, 347-354.

Haas, N., Haladyna, T. M. & Nolen, S. B. (1989). *Technical report 89-3: Standardized testing in Arizona: Interviews and written comments from teachers and administrators*. Phoenix, AZ: Arizona State University.

LeMahieu, P. & Wallace, R. (1986). Up against the wall: Psychometrics meets praxis. *Educational Measurement: Issues and Practice, 5(1)*, 12-16.

Madaus, G. F. (1987). *Testing and the curriculum*. Chestnut Hill, MA: Boston College.

Mehrens, W. A. & Kaminski, J. (1989). Methods for improving standardized test scores: Fruitful, fruitless, or fraudulent? *Educational Measurement: Issues and Practice, 8(1)*, 14-22.

Moore, W. (1991). Relationships among teacher test performance pressures, perceived testing benefits, test preparation strategies and student test performance (Doctoral dissertation, University of Kansas, 1991) *Dissertation Abstracts Internationale*.

Moore, W. (1992, Apr.). *Testing perceptions, practices, & malpractice: The impact on teachers of court-ordered achievement testing in a desegregation setting*. Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco.

Nolen, S. B., Haladyna, T. M. & Haas, N. S. (1990, Apr.). *A survey of actual and perceived uses, test preparation activities, and effects of standardized achievement tests.* Paper presented at the joint annual meetings of the American Educational Research Association and the National Council for Measurement in Education, Boston.

Popham, W. J. (1991). Appropriateness of teachers' test-preparation practices. *Educational Measurement: Issues and Practice, 10(4)*, 12-15.

Schafer, W. D. (1991). Essential assessment skills in professional education of teachers. *Educational Measurement: Issues and Practice, 10(1)*, 3-6, 12.

Schafer, W. D. & Lissitz, R. W. (1987). Measurement training for school personnel: Recommendations and reality. *Journal of Teacher Education, 38(3)*, 57-63.

Smith, M. L. & Rottenberg, C. (1991). Unintended consequences of external testing in elementary schools. *Educational Measurement: Issues and Practice, 10(4)*, 7-11.

Stiggins, R. J. (1987, Apr.). *Profiling classroom assessment environments*. Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco.

Stiggins, R. J., Conklin, N. & Faires. (1989, Mar.). *Teacher training in assessment*. Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco.

Figure 1. Proportion correctly identifying testing practices.

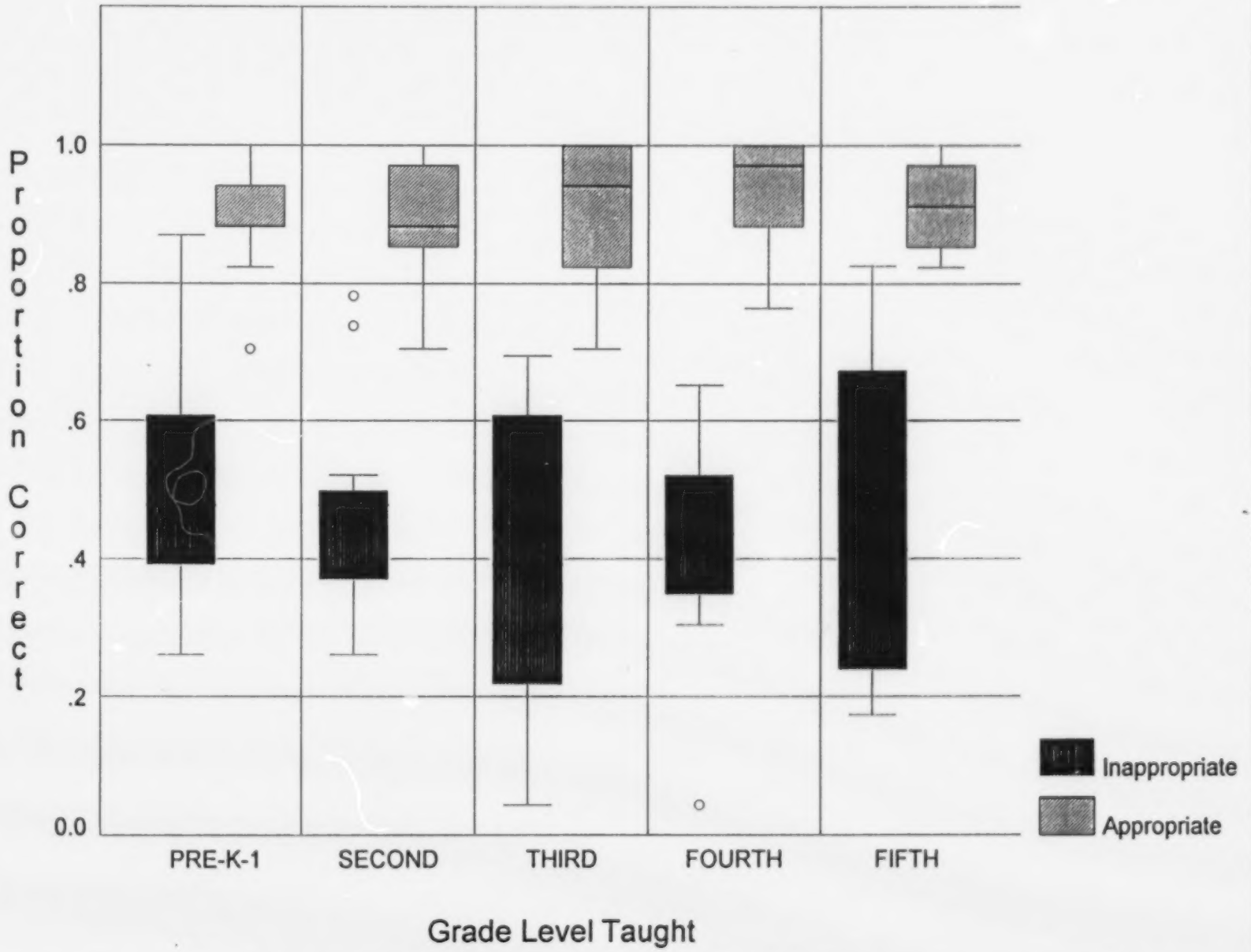

Table 1
Perceptions of Appropriateness of Testing Behaviors

	Item	Testing Behavior	Response Choice: Acceptable	Questionable	Unacceptable	Cheating
	2	Teach students how to follow test directions	1.00	.00	.00	.00
	23	Teach how to answer multiple choice questions	1.00	.00	.00	.00
	5	Discuss how to mark answer sheet correctly	.98	.02	.00	.00
	14	Discuss test-taking skills needed for ITBS	.98	.02	.00	.00
	8	Encourage good eating, sleep, and be rested for test	.96	.04	.00	.00
	19	Provide training in anxiety-reduction techniques	.92	.08	.00	.00
	4	Teach deductive reasoning skills	.86	.12	.02	.00
	30	Discuss how to re-check answers	.86	.10	.02	.02
	31	Teach clues on how to find the correct answer	.82	.12	.04	.02
	33	Create exciting classroom environment around test days by using signs, posters, and other spirit-related activities	.79	.13	.06	.02
*	35	Encourage attendance in test week and provide rewards for high attendance	.77	.17	.04	.02
	15	Teach guessing strategies	.73	.13	.15	.00
*	41	Prepare instructional objectives based on test items	.67	.25	.06	.02
*	12	Use commercial test prep package (Scoring High)	.66	.26	.02	.06
	18	Conduct special reviews or drills in prep for tests	.66	.26	.06	.02
	26	Give hints/strategies to help answer multiple choice items	.61	.16	.12	.10
*	7	Take each skill tested and direct day-to-day instruction toward these skills	.55	.31	.10	.04
*	39	Review with students skills that are on next days test	.53	.20	.08	.18
	10	Use ITBS test format for format of class tests	.49	.37	.10	.04
	16	Provide practice questions like those found on the test	.47	.18	.10	.25
	3	Provide prizes/incentives for hard work preparing	.46	.42	.13	.00
	25	Have contests before test to motivate pupils for testing	.40	.45	.15	.00
*	28	Focus instruction on extensive drill & practice on skills or items similiar to those tested	.40	.31	.27	.02
*	36	Talk to best students and encourage them to do their best on test	.40	.40	.13	.08
*	11	Assign test prep homework on weekends and vacations	.39	.35	.20	.06
*	21	Teach vocabulary words that are on current test	.35	.25	.12	.29
*	17	Give rewards for completing the test(s)	.32	.34	.30	.04
*	29	Change testing time schedule to accomodate class sche	.31	.24	.20	.24
*	20	Teach question(s) seen on past ITBS tests	.28	.30	.17	.24
*	38	Remind students to not take test too seriously	.25	.35	.23	.17
*	34	Give practice questions which are off current test with a change in the stem or distractors	.23	.17	.23	.36
*	6	Use prior yr's test questions as practice for this yr.	.22	.18	.27	.33
*	37	Give practice questions which are directly off current test	.21	.09	.04	.66
*	13	Extend testing time limits to make sure all students finish	.18	.18	.12	.51
*	22	Show past version(s) so students know what to expect	.17	.27	.27	.29
*	24	During test provide a minor hint or clue to help	.14	.09	.14	.63
*	27	Praise students who answer correct during the test	.13	.19	.27	.42
*	40	Give additional examples during testing	.13	.08	.25	.54
*	32	Encourage lower ability students to stay home on test days	.06	.00	.16	.78
*	9	Recode answer sheet because you know student just miscoded the answer	.04	.04	.13	.79

Note: All Subjects (S's) n=50; Teachers n=40; Paraprofessionals (Paras) n=10. *= unacceptable/inappropriate for ITBS.

Table 2

Perceptions of Appropriateness for Testing Behaviors: Proportion of Respondents Considering Behavior to be 'Acceptable' or 'Questionable but not Inappropriate'

	Item	Testing Behavior	Role of Participant: All S's	Teachers	Paras
	2	Teach students how to follow test directions	1.00	1.00	1.00
	5	Discuss how to mark answer sheet correctly	1.00	1.00	1.00
	8	Encourage good eating, sleep, and be rested for test	1.00	1.00	1.00
	14	Discuss test-taking skills needed for ITBS	1.00	1.00	1.00
	19	Provide training in anxiety-reduction techniques	1.00	1.00	1.00
	23	Teach how to answer multiple choice questions	1.00	1.00	1.00
	4	Teach deductive reasoning skills	.98	.98	1.00
	30	Discuss how to re-check answers	.96	.98	.88
	31	Teach clues on how to find the correct answer	.94	.93	1.00
*	35	Encourage attendance in test week and provide rewards for high attendance	.94	.95	.88
*	12	Use commercial test prep package (Scoring High)	.92	.93	.86
	18	Conduct special reviews or drills in preparation for tests	.92	.90	1.00
	33	Create exciting classroom environment around test days by using signs, posters, and other spirit-related activities	.92	.93	.86
*	41	Prepare instructional objectives based on ITBS test items	.92	.93	.88
	3	Provide prizes/incentives for hard work preparing	.88	.88	.88
*	7	Take each skill tested and direct day-to-day instruction toward these skills	.86	.85	.88
	10	Use ITBS test format for format of class tests	.86	.90	.63
	15	Teach guessing strategies	.85	.85	.88
	25	Have contests before test to motivate pupils for testing	.85	.85	.88
*	36	Talk to best students and encourage them to do their best on test	.79	.83	.63
	26	Give hints/strategies to help answer multiple choice items	.78	.76	.88
*	11	Assign test prep homework on weekends and vacations	.74	.71	.88
*	39	Review with students skills that are on the next days test	.74	.78	.50
*	28	Focus instruction on extensive drill & practice on skills or items similiar to those tested	.71	.70	.75
*	17	Give rewards for completing the test(s)	.66	.67	.63
	16	Provide practice questions like those found on the test	.65	.71	.38
*	38	Remind students to not take test too seriously	.60	.60	.63
*	20	Teach question(s) seen on past ITBS tests	.59	.63	.38
*	21	Teach vocabulary words that are on current test	.59	.61	.50
*	29	Change testing time schedule to accomodate class schedule	.56	.56	.50
*	22	Show past version(s) so students know what to expect	.44	.48	.25
*	6	Use prior yr's test questions as practice for this yr.	.41	.42	.38
*	34	Give practice questions which are off current test with a change i the stem or distractors	.40	.38	.57
*	13	Extend testing time limits to make sure all students finish	.37	.34	.50
*	27	Praise students who answer correct during the test	.31	.30	.38
*	37	Give practice questions which are directly off current test	.30	.28	.43
*	24	During test provide a minor hint or clue to help students	.22	.22	.25
*	40	Give additional examples during testing	.21	.20	.25
*	9	Recode answer sheet because you know student just miscoded the answer	.08	.05	.25
*	32	Encourage lower ability students to stay home on test days	.06	.07	.00

Note: All Subjects (S's) n=50; Teachers n=40; Paraprofessionals (Paras) n=10. *=unacceptable/inappropriate for ITBS.

Table 3
Variability of Perceptions of Appropriateness of Testing Behaviors

	Item	Testing Behavior	Measure of Variability SE	Standard Dev'n	Variance
*	37	Give practice questions which are directly off current test	.185	1.268	1.608
	16	Provide practice questions like those found on the test	.179	1.252	1.568
*	21	Teach vocabulary words that are on current test	.176	1.234	1.523
*	29	Change testing time schedule to accomodate class sche	.175	1.173	1.377
*	34	Give practice questions which are off current test with a change in the stem or distractors	.174	1.192	1.422
*	13	Extend testing time limits to make sure all students finish	.172	1.207	1.457
*	20	Teach question(s) seen on past ITBS tests	.168	1.142	1.305
*	39	Review with students skills that are on next days test	.167	1.170	1.368
*	6	Use prior yr's test questions as practice for this yr.	.165	1.158	1.342
*	24	During test provide a minor hint or clue to help	.159	1.114	1.241
*	22	Show past version(s) so students know what to expect	.155	1.075	1.156
*	27	Praise students who answer correct during the test	.153	1.062	1.127
*	40	Give additional examples during testing	.152	1.051	1.105
	26	Give hints/strategies to help answer multiple choice items	.149	1.041	1.083
*	38	Remind students to not take test too seriously	.149	1.035	1.070
*	36	Talk to best students and encourage them to do their best on test	.134	.928	.861
*	11	Assign test prep homework on weekends and vacations	.132	.922	.850
*	17	Give rewards for completing the test(s)	.130	.895	.800
*	28	Focus instruction on extensive drill & practice on skills or items similiar to those tested	.126	.871	.759
*	12	Use commercial test prep package (Scoring High)	.121	.831	.690
*	7	Take each skill tested and direct day-to-day instruction toward these skills	.119	.834	.696
	10	Use ITBS test format for format of class tests	.117	.822	.675
*	32	Encourage lower ability students to stay home on test days	.111	.779	.606
*	9	Recode answer sheet because you know student just miscoded the answer	.109	.753	.567
	15	Teach guessing strategies	.107	.739	.546
	18	Conduct special reviews or drills in prep for tests	.105	.717	.513
	25	Have contests before test to motivate pupils for testing	.103	.706	.499
*	41	Prepare instructional objectives based on test items	.103	.712	.507
	3	Provide prizes/incentives for hard work preparing	.100	.694	.482
	33	Create exciting classroom environment around test days by using signs, posters, and other spirit-related activities	.099	.689	.475
*	35	Encourage attendance in test week and provide rewards for high attendance	.095	.657	.432
	31	Teach clues on how to find the correct answer	.091	.638	.407
	30	Discuss how to re-check answers	.082	.577	.332
	4	Teach deductive reasoning skills	.060	.430	.181
	19	Provide training in anxiety-reduction techniques	.040	.277	.077
	8	Encourage good eating, sleep, and be rested for test	.029	.200	.040
	14	Discuss test-taking skills needed for ITBS	.021	.144	.021
	5	Discuss how to mark answer sheet correctly	.020	.143	.020
	2	Teach students how to follow test directions	.000	.000	.000
	23	Teach how to answer multiple choice questions	.000	.000	.000

Note: All Subjects (S's) n=50; Teachers n=40; Paraprofessionals (Paras) n=10. *= unacceptable/inappropriate for ITBS.

Table 4

Perceptions of Similarities of Testing Practices: Orthogonal Rotated Factor Loading Matrix[1]

	Item	Testing Behavior	Mode	I Inappropriate Intervention	II Testwiseness/ MDI	III Inappropriate Item Expose	IV Emphasis w/ Students	V Test-taking Skills	VI Motivational/ Incentives
•	27	Praise correct answers during the test	4	.69					
•	13	Extend testing time limits so all students can finish	4	.68					
•	32	Encourage lower ability students to stay home on test days	4	.66					
•	9	Recode answer sheet because student just miscoded answer	4	.60					
•	24	During test give minor hint or clue to help student(s)	4	.51					
•	21	Teach vocabulary words found on current test	1	.48					
	8	Encourage students to eat a good breakfast, rest, etc. for test	1	-.43					
•	40	Give additional examples during testing	4	.40					
	4	Teach deductive reasoning skills to learn to solve problems	1	-.51					
	31	Teach clues to help find correct answers to m-c items	1		.66				
	15	Teach students guessing strategies for m-c items	1		.65				
•	41	Prepare instructional objectives based on ITBS test items	1		.63				
	18	Conduct special reviews or drills in preparation for testing	1		.62				
•	28	Focus instruction on extensive drill and practice on ITBS skills	1		.57				
•	11	Assign test preparation homework on weekends and vacations	1		.51				
•	20	Teach questions seen on past ITBS tests	2		.44				
•	22	Show past versions of ITBS so know what to expect	2,3			.76			
•	34	Give practice items off current test with change in stem/dist	4			.63			
•	6	Use prior years test questions as practice for this years test	4			.56			
•	37	Give practice questions which are directly off the current test	4			.49			
•	40	Give additional examples to students during testing	4			.47			
•	36	Talk to best students and encourage them to do their best	1,2				.62		
•	39	Review with students skills that are to be tested on next days test	1				.61		
	26	Give hints/strategies to help answer m-c items	1				.52		
	16	Provide practice questions like those found on the test	1				.50		
•	38	Remind students not to take the test too seriously	2				.47		
	19	Provide training in anxiety-reduction techniques	1					.70	
	14	Discuss test-taking skills needed for ITBS	1					.64	
	30	Discuss with students how to re-check answers	1					.62	
	4	Teach deductive reasoning skills	1					.56	
	3	Provide prizes/incentives for work preparing for test	1						.77
•	17	Give rewards for completing the test(s)	2						.61
•	35	Encourage attendance in test week and reward high attendance	1						.40

Notes. [1] Six factors accounted for 47% of the common variance: I=15.5%, II= 11.2%, III=6.3%, IV=5.1%, V=4.6%, VI=4.3%.

• = Inappropriate or Unacceptable for ITBS tests. Mode 1= Acceptable, 2=Questionable but not inappropriate, 3=Unacceptable and Inappropriate, 4=Cheating. Items in matrix are those with factor loadings greater than or equal to .40.

BEST COPY AVAILABLE